AI Magazine Volume 18 Number 1 (1997) (© AAAI)

Strong AI Is Simply Silly

Selmer Bringsjord

I am deeply offended by my failure to win the Simon Newcomb Award in the past, and even now, in this victorious year, I must confess that I am greatly wounded by the sad but undeniable fact that only *one* of my rigorous attacks on "Strong" AI has managed to secure a trophy. Lest it be thought that my umbrage arises purely from matters of the heart, I hasten to reveal that my attitude is the product of a *proof*, which runs as follows. Begin by noting this principle:

> S If *p* is a fundamentally silly proposition—one which a toddler, let alone a graduate student in computer science, can see to be silly in a second—and *P* is a precise disproof of *p*, then *P* too is silly.

To see principle S in action, consider the fundamentally silly proposition *g* that your Ford is powered not by a standard combustion engine, but rather by little sedulous gremlins working feverishly inside the block. What do we make of Jones when he embarks on a protracted, precise disproof of *g*? We think that Jones' work is (at best) silly, and our disdain, *ceteris paribus*, is justified. Let *m* denote the fundamentally silly proposition that the mind is a computing machine and that cognition is computation. In light of S, it follows immediately (by universal instantiation and modus ponens) that much of my work (e.g., nearly all of *What Robots Can and Can't Be*, and, yes, the in-press argument of which Ford and Hayes are so fond) is silly. QED. So you can see why I feel so insulted. Here I sit, captain of enough silly proofs to take us, year by year, well into the next century, and I only win *once*, and only win *now*?

Of course, our attitude may change if Jones' says, "No, you don't understand! Proponents of *g* are out there; and I have set myself the task of showing that they are at bottom buffoons." We would now retract our disdain for his disproof and cheer him on. Likewise, when the uninitiated hear that there really are Strong AIniks out there in the world, they crane their necks and cup their ears to hear my arguments.

Strong AIniks will doubtless cringe upon hearing their beloved *m* classified as silly. Well, I'm sorry, but it *is*. In his inaugural writings (independent, by the way, of Turing's), Post (1936) spoke of mindless "workers," humans whose sole job was to slavishly follow explicit, excruciatingly simple instructions. This perfectly clear and seminal scheme has been supplanted by all sorts of obfuscating exotica—Turing machines, register machines, neural nets, and so on—things which facilitate formal analysis, but hide from the unsuspecting the stark, irrepressible fact that we can indeed operate as computers through and through, but doing so is for us to call upon only the dim side of our psyches. We know what it is to move the beads in an abacus; we know that such use is essentially what Post (and others) generalized in order to mathematize computation; we know that to *only* use an abacus is to be intellectually trammeled—we know all this and yet we still tolerate the Strong AInik telling us that we are always and only abaci in action. Why?

That the Strong AI is still alive may have a lot to do with its avoidance of true tests. When Kasparov sits down to face the meanest chessbot in town, he has the deck stacked against him: his play may involve super-computation, but we know that perfect chess can be played by one of Post's harebrained workers (heck, a finite state automaton is in principle sufficient), so Kasparov loses if the engineers are sufficiently clever. But when I show up to debate *m* (and related propositions), I may begin by uttering a sentence (where *M* is a fixed Turing machine the haltingness or nonhaltingness of which no Turing machine can decide, try: "It is logically possible that I decide *M*") the grasping of which may well, for all we know, require information-processing beyond the Turing Limit. At any rate, where are the word-slinging robots who will stare me down and shoot it out? Hitherto, only those humans who crave the existence of such robots show up, and they, like their progeny, are pushovers.

For the Record

- Though Strong AI is simply silly, Weak AI—the attempt to engineer systems which at least *appear* to be have minds—by my lights, is a thing of beauty. The storytelling agent BRUTUS, intended to simulate human literary creativity (and to teach us quite a bit about the nature of human creativity), marks my own latest contribution; it will debut soon (Bringsjord and Ferrucci 1997).
- While the Turing Test will soon enough be routinely passed, this will signify little, because the Turing Test is invalid (Bringsjord 1995).
- A new wave of formal arguments for Ford and Hayes to ponder can be found in Bringsjord and Zenzen (1997), in which I establish that

we are (relative to mere computation) super-minds. (So that readers really get their money's worth, I also refute Church's Thesis.) Early versions of some of these arguments can be found on my web site.

- For the ultimate in silliness, we can turn to what might be called 'Strong Sub-symbolic AI.' This view inherits all the idiocy of Strong AI *simpiciter*, because neural nets—at least ones that aren't analog and chaotic—are really just cellular automata, which are really just *k*-tape Turing machines, which are really just abaci, which takes up back to the silly notion that we don't only *use* abaci, we *are* abaci. But the sub-symbolic view adds insult to injury: it adds the impenetrability of neural nets. After all, what do we learn from PET scans about how, say, human detectives work? Zilch. Likewise, what would we learn by inspecting the innards of a connectionist system able to solve mysteries?

Bibliography

Bringsjord, S. 1997. An Argument for the Uncomputability of Infinitary Mathematical Expertise. In Expertise in Context, eds. P. Feltovich, K. Ford, and R. Hoffman. Forthcoming.

Bringsjord, S. 1995. Could, How Could We Tell If, and Why Should—Androids Have Inner Lives. In *Android Epistemology*, eds. Ken Ford, Clark Glymour, and Pat Hayes, 93–122. Cambridge, Mass.: MIT Press.

Bringsjord, S. 1992. *What Robots Can and Can't Be*. Dordrecht, The Netherlands: Kluwer.

Bringsjord, S., and Ferrucci, D. 1997. *Artificial Intelligence, Literary Creativity, and Story Generation: The State of the Art*. Hillsdale, N.J.: Lawrence Erlbaum.

Bringsjord, S., and Zenzen, M. 1997. Super-Minds: A Defense of Uncomputable Cognition. Dordrecht, The Netherlands: Kluwer. Forthcoming.

Post, E. L. 1936. Finite Combinatory Processes—Formulation 1. *Journal of Symbolic Logic* 1(3): 103–105.

Hayes and Ford Reply to Bringsjord

Professor Bringsjord is the only person to have nominated himself for our Award, and we thank him for this time-saving courtesy. He may be right to take us to task for his not having received an Award earlier. When faced with such a barrage of stuff, it is sometimes hard to discern the trees in the forest. In future, we will consider his arguments more carefully and take care to give them the consideration they deserve. However, his work does suggest the possibility that the rules committee may find it necessary to limit the number of times a single person can be given the award.

The argument sketched in his letter is not original enough, we feel. The idea that if we were computers life would seem to be indescribably tedious, since all computers are basically just abaci, reflects a distressingly familiar confusion between phenomenology and recursion theory. However, his forthcoming refutation of Church's Thesis seems likely to be a promising candidate for a Simon Newcomb Award, and we look forward to it with interest.

– Patrick Hayes & Kenneth Ford
University of West Florida

AI Magazine Errata

Sridhar Mahadevan noticed two incorrect references in his report "The National Science Foundation Workshop on Reinforcement Learning" (by Sridhar Mahadevan and Leslie Pack Kaelbling, *AI Magazine* 17(4): 89-97.) In one of the references, the author's names were reversed (Russell and Parr should be Parr and Russell). In the other reference some coauthors were unintentionally omitted. The corrected references are:

Boutilier, C.; Dearden, R.; and Goldszmidt, M. 1995. Exploiting Structure in Policy Construction. In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, 1104-1111. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Parr, R., and Russell, S. 1995. Approximating Optimal Policies for Partially Observable Stochastic Domains. In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, 1088-1094. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.